

# Evaluating Model Accuracy for Model-Based Reasoning

Steve Chien and Joseph Roden  
Jet Propulsion Laboratory, California Institute of Technology  
Pasadena, CA

## Abstract

Model-based reasoning has been proposed as a general methodology for such diverse tasks as monitoring, diagnosis, control, and design. In this approach, a behavioral model mimicking the structure of the target system is used to reason about expected performance of the target system. However, most such work does not explicitly account for inaccuracies in the model.

This paper describes an approach to automatically assessing the accuracy of various components of a model. In this approach, actual data from operation of the target system is used to drive statistical measures to evaluate the prediction accuracy of various portions of the model. We describe how these statistical measures of model accuracy can be used in model-based reasoning for monitoring and design. We then describe application of these techniques to monitoring and design of the water recovery system of the Environmental Control and Life Support System (ECLSS) of Space Station Freedom.

**Keywords:** model-based monitoring, diagnosis, control and design, validation of knowledge-based systems, model-based simulation

## 1. Introduction

Model based reasoning has been advocated as a general approach to a wide variety of tasks such as monitoring [Doyle et al. 89, Doyle et al. 91, Dvorak & Kuipers 89], diagnosis and interpretation [Davis and Hamscher 88], control [Scarl et al. 88], and design [Chien et al. 91a, Chien et al. 91b, Bose & Rajamoney 91]. However, despite this strong effort, comparatively little work has focused upon using actual data on model

performance to characterize how well a model captures the behavior of the target system.

This paper describes a statistical approach to measuring the prediction error of a model based upon an analysis of model prediction performance on actual data. This analysis produces a statistical model of expected model prediction error. This model of the model error is then used in the model-based reasoning tasks of monitoring and design.

The next section of this paper describes how the statistical techniques are used to create a model of the error and how this model of the error can be used to calculate confidence intervals. The following section describes how this confidence interval information can be used in model-based monitoring and design tasks. This section also describes several applications of this error model to monitoring and design of the Environmental Control and Life Support System for Space Station Freedom. The discussion section of this paper focuses upon ongoing work to increase the accuracy of the error models by applying machine learning techniques to learn error models.

## 2. Evaluating Model Accuracy

Model-based reasoning uses a model of a system to predict the behavior of the system under the conditions included in the scope of the model. It is useful for applications using a model to know how accurately the model predicts the behavior of a system being modeled. Instead of simply predicting that a measure will take on some value, it is more useful to state how confident the model is in predicting that value.

---

Authors' Address: JPL, M/S 525-3660,  
4800 Oak Grove Dr., Pasadena, CA 91109-8099  
Internet: {chien,roden}@aig.jpl.nasa.gov

Model accuracy can be evaluated by comparing model behavior to observed system behavior. Through analysis of errors in predicting system behavior, we can estimate the amount of error we expect a model to produce. Data obtained by performing model evaluation studies can provide a basis on which to model the errors a model produces.

Model error ( $\Delta$ ) is defined as the difference of the model predicted value ( $m$ ) based on previous observed system values and the current observed system value ( $o$ ) for a given system state:

$$(1) \quad \Delta = m - o$$

For our applications, the time step is relatively constant. Thus the model is making a prediction of the  $i$ -th time step from the data of the  $(i-1)$ th time step. In Figure 1, the model error is the difference between the model predicted value and the observed system value over time.

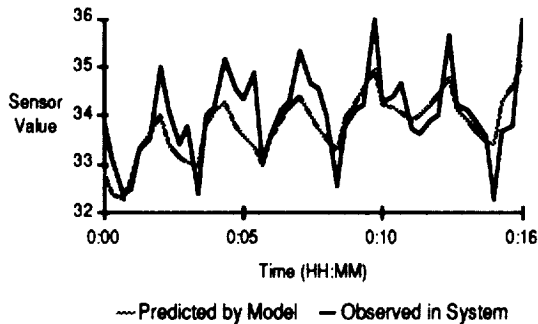


Fig 1. Graph showing difference between model predicted value and observed system value over time

For a given operating mode of a system, some number ( $n$ ) of observations of one time step model predictions are taken, and the error computed for each. This results in a sampled model error distribution of  $\Delta_1, \Delta_2, \dots, \Delta_n$ . From this distribution we develop a general model of the error. We observed our samples to be approximately normally distributed. Figure 2 shows a histogram of model error for a particular sensor. Given this sampled error distribution, we estimate

that the true model error is normally distributed with mean  $\bar{X}$  and variance  $s^2$ .

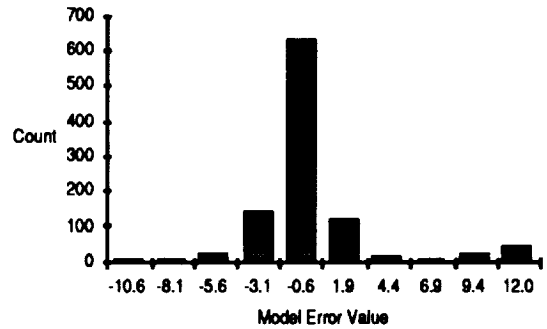


Fig 2. Histogram showing model error distribution for sensor KP02

For a given operating mode of a system, we evaluate model accuracy by determining the probability ( $p$ ) that the true system measurement ( $t$ ) will take on a value within a range of values ( $m-e, m+e$ ) around the model predicted value ( $m$ ). This probability can be determined given the distribution of the model error. Alternatively, by setting an acceptable probability bound, we can determine the range of values that the true system will take on with that probability. Equation 2 shows how this probability is determined given  $\Phi$ , the cumulative distribution function of the standard normal distribution.

$$(2) \quad P(m-e \leq t \leq m+e) =$$

$$\Phi\left(\frac{e - \bar{X}}{s}\right) - \Phi\left(\frac{-e - \bar{X}}{s}\right)$$

Using the probability table for the standard normal distribution, this measure quantifies the accuracy with which a model predicts the true system measurement.

### 3. Applying the Error Model in Model-based Reasoning

This section describes two applications of confidence intervals to model-based reasoning: model-based monitoring as applied to discrepancy detection and model-based prediction for design.

### 3.1 Application to Model-based Monitoring

One application of model accuracy is in model-based monitoring [Doyle et al. 89, Dvorak & Kuipers 89]. In model-based monitoring, a model of the target system is used to predict sensor values. Deviations from the predicted values are indications of abnormal behavior and are thus indicative of sensors which should be reported to operators. However, if portions of the model in certain operating modes are inherently inaccurate because of noise or poor understanding or predictability of the occurring phenomenon, the strength of the model's predictions should be correspondingly reduced.

One method to account for model inaccuracy in model-based monitoring is use of a running average of model/actual deviation [Doyle et al. 91]. In this approach, a running average of the deviation between model predicted and actual values is maintained. By tracking the current deviation minus the running average of the deviation, the current deviation can be ignored in cases where the model has not been tracking the system behavior accurately.

While the deviation from running average deviation measures the recent performance of the model, statistical measures over all available historical data provide a measure of past historical performance. With a confidence interval capability as described in the previous section, a more direct approach to calibrating deviation scores according to model accuracy can be applied. Specifically, the statistical model of the prediction error of the model can be used to generate a measure of the unusualness of a deviation of the model from the observed value.

For example, consider the following example from our ECLSS monitoring application. Using the techniques described in the previous section, a sensor KP02 is measured to have a error with a measured distribution of mean -0.19 and standard deviation 1.89 in system operating mode PROCESS. If we observe a discrepancy of 3 PSIG between the observed and predicted values, we can now use the equations shown

in Section 2 to produce a confidence rating of 0.89 that the model is within 3 PSIG of the actual. Thus high confidence values for the error being less than the current deviation indicate unusual deviations.

As a second application to model-based monitoring, consider the case where a model historically predicts well but has recently been predicting poorly. This may indicate a persistent unexplained phenomenon affecting the sensor or portion of the system in question. Such a situation could be detected by determining if the running average of the deviation is at a level which is relatively unusual given the error model (e.g. for a running average deviation  $e$ ,  $P(-e \leq \Delta \leq e)$  is high). This provides a measure of the unusualness of the running average of the deviation. Note that this monitoring measure and the previous one are complementary. The unusualness of the current deviation catches quickly developing departures from normal operations, but is susceptible to random noise. The unusualness of the running average of the deviation is not susceptible to random noise, but takes longer to manifest and inform.

### 3.2 Application to Model-based Prediction for Design

Another application of model-based reasoning is in evaluating sensor placements for assistance in the design process. Specifically, we have been working on approaches to evaluate sensor placements with respect to a diagnosability criterion [Chien et al. 91]. In this approach a model of the target system is used to determine how specific proposed sensors would report altered scores in the event of a fault occurrence. More specifically, we evaluate how well a sensor can distinguish between classes of states with respect to three criteria. For the purposes of fault detection, the relevant distinction is between faulted and non-faulted states. For the purposes of fault isolation, the relevant distinction is between faulted states.

Towards evaluating diagnosability, we have developed three measures. First, *Discriminability* measures how much of a

divergence the model predicts would occur in comparing between the two states. Second, *Accuracy* measures the confidence in the model's prediction of the expected divergence. Third, *Timeliness* measures the time lag between the occurrence of the fault and the discrimination detected by the sensor.

Using the measure for model error we have described in this paper, we can formulate the confidence that the predicted divergence would be predicted and the actual value not deviate as a probability. The lower the probability of this occurrence, which represents the model predicting a change that does not occur, the more likely the sensor will be able to perform the discrimination.

### 3.3 Examples from Application to the ECLSS Testbed

Our sensor placement approach is being tested upon the water reclamation subsystem of the Environmental Control and Life Support System (ECLSS) for Space Station Freedom. A model describing the behavior of the Multifiltration Subsystem (MF) in terms of fluid flow and heat transfer has been constructed. This model was developed via a combination of study of design documentation (i.e., schematics, etc.) and consultation with domain experts (e.g. the

operators of the testbed). This model has been validated by comparison against actual data from the subsystem testbed undergoing evaluation at the Marshall Space Flight Center in Huntsville, Alabama. We also have constructed models of the Vapor Compression and Distillation (VCD) and Volatile Removal Assembly (VRA) subsystems of SSF ECLSS. Together, these models represent coverage of virtually the entire water-side of SSF ECLSS. We are also in the process of extending our model to cover ECLSS air-side subsystems.

Figure 3 below shows the ECLSS multifiltration subsystem. In this subsystem, the water first passes through a pump at the inlet to the MF system. Next, the water passes through a coarse filter before entering the sterilization loop. In the sterilization loop the water is heated in the regenerative heat exchanger and then by the in-line heater after point 3. Within the sterilizer reservoir, the temperature of the water is maintained at 250° F for several minutes. In the second portion of the subsystem, the water passes through a set of unibed filters designed to remove particulate contaminants from the water. Possible sensor types are flow rate, water pressure, and temperature. Possible sensor locations are indicated by ovals in Figure 3.

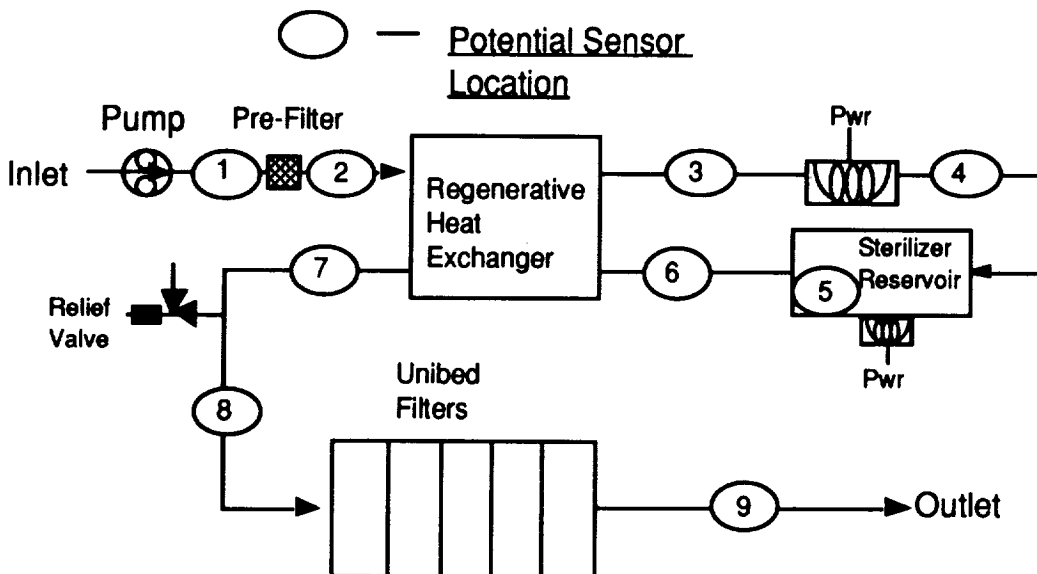


Fig 3. The Multifiltration Subsystem

In model-based monitoring, the empirically derived model accuracy scores impact monitoring in the following way. The process mode model of the conductivity sensor KT02 at point 4 exhibits poor model accuracy (empirically derived mean of 2.15 and standard deviation of 27.24). Thus, relatively large deviations from model predicted values such as 4, with a confidence rating of 0.52, do not cause the sensor score to be brought to the attention of the operator. However, the process mode model of sensor KP02 is more accurate (mean -0.19 and standard deviation 1.89) so that relatively small deviations on the order of 4, with a confidence rating of 0.97, cause the sensor to be flagged and the sensor value to be brought to the attention of the operator.

In model-based diagnosability assessment, again the model accuracy figures heavily in evaluating certain sensors. For example, one possible fault is unibed loading, which occurs when particulate matter gets caught in the unibed filters. This fault has several effects. First, a pressure drop would occur, causing a lower pressure at location 9. Second, unibed performance would decrease, resulting in an increase in conductivity downstream from the unibeds. Third, if loading is significant, flow in the entire subsystem may decrease. Again, because the conductivity models are not very accurate, the *Accuracy* measure of the diagnosability evaluation would score the pressure sensor placement higher than the conductivity sensor placement for fault detection of this fault.

#### 4. Discussion

This work is preliminary; there are a number of outstanding issues. One issue is the selection of a normal distribution to model the error. Other possible distributions may model the error more accurately. A measure of how well the derived error model matches the observed distribution would be useful in assessing the degree of confidence in the error model.

Another issue is our choice to model the error in absolute terms rather than as a

percentage bound based upon the current model prediction (e.g. 4 PSIG  $\pm$  5% rather than 4 PSIG  $\pm$  .20). This has ramifications if the model error tends to increase as a function of the model predicted value. A cursory analysis indicates that in general, in our domain, the error is not strongly correlated with the model predicted value so that modelling the absolute error seems reasonable.

We also model the model error independent of potentially relevant factors such as other causally related model predicted values. For example, the model may use an equation to derive a temperature in the MF subsystem that is accurate only in cases where the water pressure is high. One extension of our work focuses upon using machine learning techniques to determine what other potentially relevant factors would be good indicators of model accuracy. In this work we are investigating applying GID3\* to learn a model accuracy function for a sensor S based upon model predicted value for S and other sensors.

Another outstanding issue is that of dealing with variable time steps. The accuracy of the model's predictions clearly depends upon how far into the future the model is required to make predictions. Currently, our model of the model prediction error does not account for this variable.

#### 5. Conclusion

This paper has described an approach to evaluating the accuracy of a model's predictions. This approach uses statistical methods to develop a model of expected error in model predictions. This paper has also described how this statistical measure for model error can be used in two model-based reasoning tasks: model-based monitoring and model-based reasoning for evaluating sensor placements. By application of our derived measure for model accuracy, the degree of accuracy of the model of the target system can be accounted for to increase the usefulness of model-based reasoning in both monitoring and evaluation of sensor placements.

## Acknowledgements

This work was performed by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

## References

- [Bose and Rajamoney 91] P. Bose and S. Rajamoney, "An Approach Based on First Principles For The Design of Continuous Devices," *Proceedings of the 1991 Workshop on Model-based Reasoning*, Anaheim, CA, 1991.
- [Chien et al. 91a] S. A. Chien, R. J. Doyle, and L. S. Homem de Mello, "A Model-based Reasoning Approach to Sensor Placement," *Proceedings of the 1991 Workshop on Model-based Reasoning*, Anaheim, CA, 1991.
- [Chien et al. 91b] S. A. Chien, R. J. Doyle, and N. Rouquette, "A Model-based Reasoning Approach to Sensor Placement for Diagnosability," *Proceedings of the Second International Workshop on the Principles of Diagnosis*, Milan, Italy, 1991.
- [Doyle et al. 89] R. J. Doyle, S. M. Sellers, and D. J. Atkinson, "A Focused Context-sensitive Approach to Monitoring," *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, Detroit, MI, 1989.
- [Doyle et al. 91] R. J. Doyle, U. M. Fayyad, D. Berleant, L. K. Charest, L. S. Homem de Mello, H.J. Porta, and M.D. Wiesmeyer, "Sensor Selection in Complex Systems Monitoring Using Information Quantification and Causal Reasoning," in *Recent Advances in Qualitative Physics*, B. Faltings and P. Struss (eds.), MIT Press, 1991.
- [Dvorak and Kuipers 89] D. Dvorak and B. Kuipers, "Model-based Monitoring of Dynamic Systems," *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, Detroit, MI, 1989.
- [Davis and Hamscher 88] R. Davis and W. C. Hamscher, "Model-based Reasoning: Troubleshooting," In H. E. Shrobe, editor, *Exploring Artificial Intelligence: Survey Talks from the National Conference on Artificial Intelligence*, Morgan Kaufman, San Mateo, CA 1988.
- [Scarl et al., 88] E. A. Scarl, J. R. Jameison, and E. New, "Deriving Fault Location and Control from a Functional Model," *Proceedings of the Third IEEE Symposium on Intelligent Control*, Arlington, VA, 1988.